Relationships of p16 immunohistochemistry and other biomarkers with diagnoses of cervical abnormalities: implications for LAST terminology

Philip E. Castle, PhD, MPH; Rachael Adcock, BS; Jack Cuzick, PhD; Nicolas Wentzensen, MD, PhD; Norah E. Torrez-Martinez, BS; Salina M. Torres, PhD, MPH; Mark H. Stoler, MD; Brigitte M. Ronnett, MD; Nancy E. Joste, MD; Teresa M. Darragh, MD; Patti E. Gravitt, PhD; Mark Schiffman, MD; William C. Hunt, MS; Walter K. Kinney, MD; Cosette M. Wheeler, PhD

on behalf of the New Mexico HPV Pap Registry Steering Committee and for the p16 IHC Study Panel

Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY, USA (Dr. Castle); Wolfson Institute of Preventive Medicine, Queen Mary University of London, London, UK (Ms. Adcock, Dr. Cuzick); Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Rockville, MD, USA (Dr Wentzensen, Dr. Schiffman); Department of Pathology, University of New Mexico Cancer Center, Albuquerque, NM, USA (Ms. Torrez-Martinez, Dr. Torres, Dr. Joste, Dr. Gravitt ,Mr. Hunt, Dr. Wheeler); Department of Pathology, University of Virginia Health System, Charlottesville, VA, USA (Dr Stoler); Department of Pathology, Johns Hopkins University, Baltimore, MD, USA (Dr Ronnett); Department of Pathology, University of California, San Francisco, CA, USA (Dr Darragh); Sacramento, CA, USA (Dr Kinney); The New Mexico HPV Pap Registry Steering Committee Members are listed on Page 24 and 25; The p16 IHC Study Panel Members are listed on Page 25

Dr. Gravitt is currently located at the Department of Epidemiology and Public Health, University of Maryland School of Medicine, Baltimore, MD, USA

Corresponding Author

Cosette M. Wheeler, PhD

Department of Pathology
University of New Mexico Cancer Center
MSC02-1670 House of Prevention Epidemiology
1 University of New Mexico
Albuquerque, NM 87131
Phone (505)269-5817 Fax (505)277-0265 cwheeler@salud.unm.edu

Disclosures: PEC has received HPV tests and assays for research at a reduced or no cost from Roche, Cepheid, Becton Dickinson, and Arbor Vita Corporation. JC has received funds from grants, cooperative agreements or subcontracts related to cervical screening and triage through his institution. JC reports grants and personal fees from Qiagen, grants from Hologic, grants and personal fees from Becton Dickinson, grants and personal fees from Genera Biosystems, grants from Gene First, grants from Trovagene, outside the submitted work. NW is an employee of the National Cancer Institute, which has received HPV tests at reduced cost from BD and Roche. MHS is a consultant and has received fees for clinical trial development and as an expert pathologist in clinical trials for Merck, Roche/Ventana Medical Systems, Becton Dickinson (BD), Hologic /Gen Probe, Cepheid & Inovio Pharmaceuticals and was Co-Lead of Work Group 4 and an author of the College of American Pathologists (CAP)-American Society of Colposcopy and Cervical Pathology (ASCCP) LAST Project. BMR received consulting fees from Merck as a member of pathology adjudication panel for HPV vaccine trials. TMD has received consultant fees from BD, Roche, Antiva and Thevax; research supplies for anal cytology from Hologic and was lead author on the CAP-ASCCP LAST Project. CMW has received funds from grants and cooperative agreements related to cervical screening and triage through her institution. CMW reports receiving reagents for HPV genotyping from Roche and reagents and equipment from Roche/Ventana Medical Systems through her institution and outside of the submitted work, personal fees from BD. The other authors have no relevant financial interest in the products or companies described in this article.

ABSTRACT

Context. Lower Anogenital Squamous Terminology (LAST) standardization recommended $p16^{INK4a}$ immunohistochemistry (p16 IHC) on biopsies diagnosed morphologically as cervical intraepithelial neoplasia (CIN) grade 2 (CIN2) to classify them as low-grade or high-grade squamous intraepithelial lesions (HSIL).

Objective. To describe the relationships of p16 IHC and other biomarkers associated with cervical-cancer risk with biopsies diagnoses.

Design. A state-wide, stratified sample of cervical biopsies diagnosed by the community pathologists (CP), including 1,512 CIN2, underwent a consensus, expert pathologists panel (EP) review (without p16 IHC results), p16 IHC interpreted by a third pathology group, and HPV genotyping, results of which were grouped hierarchically according to cancer risk. Antecedent cytologic interpretations were also available.

Results. Biopsies were more likely to test p16 IHC positive with increasing severity of CP diagnoses, overall (P_{trend}<.001) and within each HPV risk group (P_{trend}≤.001). All abnormal grades of CP-diagnosed biopsies were more likely to test p16 IHC positive with a higher HPV risk group (P_{trend}<.001), and testing p16 IHC positive was associated with higher HPV risk group than testing p16 IHC negative for each grade of CP-diagnosed biopsies (*P*<.001). p16 IHC-positive, CP-diagnosed CIN2 biopsies were less likely than CP-diagnosed CIN3 biopsies to test HPV16 positive, have an antecedent HSIL+ cytology, or to be diagnosed as CIN3+ by the EP (*P*<.001 for all). p16 IHC-positive, CP-diagnosed CIN1 biopsies had lower HPV risk groups than p16 IHC-negative, CP-diagnosed CIN2 biopsies (*P*<.001).

Conclusions. p16 IHC-positive, CP-diagnosed CIN2 appears to be lower cancer risk than CP-diagnosed CIN3. LAST classification of “HSIL” diagnosis, which includes p16 IHC-positive

CIN2, should annotate the morphologic diagnosis (CIN2 or CIN3) to inform all management decisions, which is especially important for young (<30 years) women diagnosed with CIN2 for whom surveillance rather than treatment is recommended.

INTRODUCTION

Persistent cervical infections by 12-15 high-risk human papillomavirus (HPV) genotypes cause nearly all cervical cancers[1] and most of the immediate precursor cervical abnormalities, including cervical intraepithelial neoplasia (CIN) grade 2 (CIN2), grade 3 (CIN3), and adenocarcinoma *in situ* (AIS). HPV16 and HPV18 are the most carcinogenic HPV genotypes, with HPV16 causing approximately 50-60% of cervical cancers and HPV18 causing 10-15% of cervical cancers.[2] The other 10-13 HPV types cause the remaining 25-40% of cervical cancers.[2] With increasing severity of the cervical abnormality, attributable fractions due to HPV16 and HPV18 increase while those due to other types concomitantly decrease.[3]

CIN2 has been the threshold for cervical treatment, by either excision or ablative treatment.[4] However, recently there has been an increasing recognition that the H&E diagnosis of CIN2 is an equivocal diagnosis with significant inter-observer variability and likely represents an admixture of (misclassified) HPV infection/CIN1 and precancer (CIN3)[5] rather than a biological intermediate step in the progression from CIN1 to CIN3 as was originally thought.[6] The uncertainty of the meaning of this diagnosis is perhaps reflected in its poor diagnostic reproducibility between pathologists.[7-11] Because CIN2 likely has overall low immediate potential to become invasive cancer, frequently regresses especially in young women (aged <30 years)[12], and excisional treatment is possibly associated with an increased risk of preterm delivery[13, 14], current management guidelines in the United States (U.S.) recommend "wait and watch" rather than treatment for CIN2 diagnosed in young women (aged <30 years) of reproductive potential when the squamocolumnar junction can be visualized in its entirety.[4]

There has been great interest in using adjunctive biomarkers to improve the classification and reliability of histopathologic diagnoses, based on hematoxylin and eosin (H&E) staining, of cervical abnormalities, especially to reduce the over-diagnosis of CIN2 on H&E, clarify the clinical significance of CIN2 (i.e., distinguish between benign CIN2 diagnoses potentially destined to regress or not progress from CIN2 diagnoses that reflect the presence of high-grade cervical abnormalities that, for safety against cancer, should be treated to reduce the risk of cancer development.). Some of the biomarkers investigated for clarifying the meaning of an H&E diagnosis of CIN2 on biopsy include (but are not limited to) HPV16[15], HPV L1[16, 17], Ki-67[7, 16], E4[18, 19], and $p16^{INK4a}$ (p16)[7, 11, 16] detection.

Immunohistochemistry (IHC) for *in situ* detection of p16 (p16 IHC) has emerged as an adjunctive biomarker to aid in the diagnosis of cervical abnormalities. p16 IHC has been shown to be sensitive for CIN2 and CIN3[7, 14] and its interpretation is much more reliable/reproducible than morphology based on H&E staining alone.[11, 20, 21] Recommendations from the Lower Anogenital Squamous Terminology (LAST) Standardization Project include the use of p16 IHC in the following specific circumstance[22]: "If the pathologist is entertaining an H&E morphologic interpretation of –IN 2 (under the old terminology, which is a biologically equivocal lesion falling between the morphologic changes of HPV infection [low-grade lesion] and high-grade cervical abnormalities), p16 IHC is recommended to help clarify the situation. Strong and diffuse block-positive p16 IHC results support a categorization of precancer. Negative or non-block-positive staining strongly favors an interpretation of low-grade disease or a non-HPV-associated pathology." LAST recommended a switch from the three-tier categorization, CIN1, CIN2, and CIN3, to a two-tier system of categorization of low-grade squamous intraepithelial lesion (LSIL),

which includes CIN1 and p16 IHC-negative CIN2, and high-grade squamous intraepithelial lesion (HSIL), which includes CIN3 and p16 IHC-positive CIN2.

However, the question remains about whether p16 IHC distinguishes between benign HPV infection and clinically significant CIN2 i.e., those that have or will develop invasive potential thereby representing a high-grade cervical abnormality. Obviously, it is not logistically or ethically possible to follow a cohort of women diagnosed with CIN2 to see who develops cervical cancer to answer this question, as was done tragically with CIN3/carcinoma *in situ*.[23] The subsequent diagnosis of CIN3 in follow-up of CIN2 cases may not be true progression but rather a correction of a previously misclassified CIN2 diagnosis and sampling errors including missed CIN3 at colposcopy.

To better understand the cervical-cancer risk stratification achieved by p16 IHC for routine diagnoses (community pathology [CP]) of CIN2 as well as other diagnosis, we conducted a large U.S. population-based study of p16 IHC and its relationship to other biomarkers of cervical-cancer risk, including an expert panel (EP) consensus review that has been shown to improve the certainty of high-grade cervical abnormalities and therefore the association with HPV[24], tissue HPV genotyping, and antecedent cytology result. Increasing severity of histologic diagnosis rendered by an EP review of a CP diagnosis of CIN2 is associated with a CIN3 diagnosis on tissue from an excision procedure.[25] The percent positive for the highest risk HPV genotypes, especially HPV16, increases with the severity of cervical diagnosis[3, 25], and the HPV genotype(s) detected in the diagnostic tissue generally is considered the cause of the cervical abnormality. Cytologic interpretations of high-grade squamous intraepithelial lesion (HSIL) or more severe (HSIL+) are more strongly associated with histologically confirmed CIN3 and cancer than less severe cytologic interpretations, and antecedent HSIL often proceeds rare cases

of invasive cervical cancer in the follow-up of women diagnosed with CIN2 and under surveillance (vs. immediate treatment).[26, 27] HSIL cytology is of sufficient clinical concern that treatment is considered acceptable even without histologic confirmation of CIN2 or more severe diagnoses (CIN2+) on biopsy.[4]

Our main goal was to assess whether biopsy diagnosed as CIN2 by morphology and tested positive for p16 by IHC was similar enough to biopsy diagnosed CIN3 in the distribution of these other biomarkers of cervical cancer risk such that making a distinction between the two would be unnecessary i.e., calling both HSIL without annotating the morphologic diagnosis of CIN2 or CIN3.

METHODS

Cervical biopsies used in the current study were part of a previous population-based study.[10] The biopsy with the most severe diagnosis of individual women diagnosed in the period of 2006-2009 was used. Of the 21,297 women diagnosed in laboratories serving New Mexico's residents during the study period, a stratified sample of 6,272 women was chosen to over-represent CIN2 and CIN3 for additional characterization. This sample included 90.1% of all CIN2+ diagnosed, which represented all adequate CIN2+ biopsies that could be found, and random samples of 17.7% of all CIN1 and 6.3% of all negative histology biopsies diagnosed during that period.

Laboratory Testing

A "sandwich" technique was employed to enable histopathologic review of tissue sections flanking the sections subjected to HPV genotyping and p16 IHC as follows: One four micron (4 μM) section was obtained for H&E staining, two 4 μM sections for HPV genotyping were collected into o-ringed microfuge tubes, a second 4 μM section was obtained for H&E staining, and then 4μM section(s) adjacent to this second H&E were obtained for biomarker staining including p16 IHC with sections collected onto Fisherbrand Superfrost Plus glass slides.

Selection of the cases for the current study was limited to those in which 5 or more unstained slides were available to allow for potential unsatisfactory slides and the opportunity for further tests on the same subset. This resulted in a group of 4,359 cases from which 4,100 biopsies from different women were selected randomly to create 41 sets of 100 slides, each of which was reviewed by one of 41 different pathologists participating in this study as volunteers (p16 IHC Study Group). Patient age and, when available, referral cytology (3,563, 86.9% of

cases) were also provided to the pathologists. p16 IHC was performed, as described below, on an unstained slide adjacent to the H&E stained slide used for diagnosis from each case. Each set contained similar proportions of each diagnostic outcome.

p16 IHC. Formalin-fixed, paraffin-embedded (FFPE) cervical tissue sections were stained using one of two methods; manually using the CINtec Histology Kit (Roche mtm laboratories) or using the automated BenchMark instrument platform (Ventana/Roche). Briefly, de-paraffinization was performed by baking the slides at 65°C for 45 minutes followed by rehydration of the tissues in xylene and graded alcohol baths (95%, 70%, and 50%). Optimized epitope retrieval for archival tissues was performed at 95°C for 45 minutes in epitope retrieval solution. Epitope retrieval slides were either transferred to the BenchMark XT or manually stained. The p16 IHC staining and visualization procedures followed the BenchMark p16 IHC protocol or the CINtec Histology Kit protocol specified by the manufacturer.

HPV genotype detection in FFPE tissues. Methods for HPV genotyping of the FFPE tissue sections were previously reported and are summarized here.[28] FFPE tissue sections were digested in a protein K (PK) digestion buffer at 65°C for 4 hours followed by overnight at 37°C. Prior to polymerase chain reaction (PCR)-based HPV genotyping, digested FFPE tissue was heated at 95°C for 15 minutes to inactivate the PK and centrifuged briefly at 13,000 × g to remove undigested material, and the supernatant (aqueous digest) was decanted and stored at -80°C until tested.

Two and five μL (for two separate genotyping determinations) of the aqueous digest from each tissue specimen were used for genotyping with the LINEAR ARRAY HPV Genotyping Test (HPV LA; Roche Diagnostics, Indianapolis, Indiana USA), a qualitative HPV genotyping

test for 37 HPV genotypes.[29-31] Using the Roche LA HPV detection kit, hybridizations were automated using Tecan ProfiBlot-48 robots (Tecan, Grödig, Austria) as previously described.[32] The Roche LA HPV Genotyping Test detects 13 high- and 24 low-risk HPV types. HPV52 is not determined directly by a type-specific probe but by inference as previously described.[28, 31] Two independent readers interpreted the presence of HPV genotypes using a reference template and any differences between the two readers were adjudicated by a third independent reader. The adjudicated result was taken as the final interpretation.

Pathology Reviews

EP Reviews. EP pathologists rendered an adjudicated consensus diagnosis of these biopsies.[10] The EP diagnosis review was based only on an H&E staining of a new section and masked to any other data including p16 IHC results when available.

Volunteer Pathologist (VP) Reviews. 41 pathologists (p16 IHC Study Panel Group; here referred to as VP) practicing throughout the United States (US) and Canada agreed to review the H&E and p16 IHC for 100 cervical biopsies each. Recruitment of VP pathologists was either via direct invitation for College of American Pathologist committee members or through general advertisement (e.g. flyers) at professional society meetings.

VPs were given a two-page instructions sheet developed by Ventana to guide their interpretation of p16 IHC. Diffuse p16 IHC staining was considered positive ("A continuous staining of cells of the basal and parabasal cell layers of the cervical squamous epithelium, with or without staining of the intermediate or intermediate to superficial cell layers."). p16 IHC resulting in a focal ("Either a staining of isolated cells or small cell clusters; i.e., a non-continuous staining, particularly not of the basal and parabasal cells.") or negative ("the p16

stained slide shows no staining reaction") staining pattern were considered negative. Only the p16 IHC interpretations of the VP were used in theses analyses.

The p16 IHC interpretations of the VP were used whereas morphologic diagnoses of the VP, though performed for studies of p16 utilization, were not included in these analyses. Use of p16 interpretations rendered by the independent VP was preferred to reduce inherent interpretation biases when comparing community and EP diagnoses.

Analysis

The primary aim of the analysis was to examine whether the CP-diagnosed CIN2 that subsequently tested p16-IHC positive was the risk equivalent to the CP-diagnosed CIN3, which is routinely treated in clinical practice. Thus, the CP diagnoses, with and without stratification by p16 IHC as read by the VP, were compared to the EP diagnoses (masked to p16 IHC results), tissue HPV genotyping, and antecedent cytologic interpretations.

Of the statewide sample of 21,297 cervical biopsies, there were 21,187 cervical biopsy tissues after excluding AIS, adenocarcinoma, and cancers other than squamous cell carcinoma (SCC) histology. A consort diagram of the biopsies included in this analysis is shown in **Figure 1**. Sampling fractions of the statewide population for each grade of biopsy diagnosis by the CP used in these analyses are shown in **Supplemental Table 1** for reference. For simplicity, these analyses did not correct for sampling fractions (of the CP diagnoses) (except where otherwise noted) as all CIN2+ biopsies that could be located were collected and were assumed to be representative i.e., cases were missing at random. A random sample of CP-diagnosed negative and CIN1 histology were included from the entire state of New Mexico.

Of the 4,100 samples tested by p16 IHC, 65 were excluded (from above) from this analysis because they were diagnosed by CP as glandular disease, adenocarcinoma *in situ* or adenocarcinoma, which are not included in the LAST recommendations. An additional 25 were classified as "Technically Unsatisfactory" by any (CP, EP, and/or VP) pathology review, thus excluded. As noted, the distribution of HPV genotypes causing squamous cell carcinoma (SCC) were included for reference under the assumption that the more closely the profile of biomarkers for a precursor diagnosis resembled SCC, the better proxy it was for cancer risk i.e., high-grade cervical abnormalities with invasive potential. After the aforementioned exclusions, the resulting sample size was 4,010.

HPV genotype results were categorized hierarchically according to their *a priori* cancer risk[2, 3, 33] for simplicity of presentation and to account for detection of multiple HPV genotypes: 1) HPV16 positive; 2) HPV16 negative and HPV18 and/or HPV45 (HPV18/45) positive, 3) HPV16, 18, and 45 negative and HPV31, 33, 35, 39, 51, 52, 56, 58, 59, and/or 68 positive (other high-risk HPV) positive, 4) negative for all high-risk HPV types and HPV26, 53, 66, 67, 70, 73, and/or 82 (intermediate risk HPV) positive, 5) high- or intermediate-risk HPV negative and HPV6, 11, 40, 42, 54, 55, 61, 62, 64, 69, 71, 72, 81, 82v, 83, 84, and/or 89 (low-risk HPV) positive, or 6) HPV negative for all measured types. A shift in a distribution to higher- or lower-HPV risk groups was considered to be related to a greater or lesser risk of cervical cancer, respectively.

The proportion and binomial exact 95% confidence interval (95%CI) of p16 IHC positive by grade of biopsy diagnosis by CP was calculated. A non-parametric test of trend[34] or Trend test using weighted logistic regression model was used to test for trends in p16-IHC or HPV risk groups within or between diagnostic categories by the CP or EP. Trends in the percent p16-IHC

positive, HPV16 positive, and with an antecedent HSIL+ cytology, according to The Bethesda System for cytologic classification[35], for paired CP and EP diagnoses were calculated. The percent p16-IHC positive, HPV16 positive, with an antecedent HSIL+ cytology, and with cervical intraepithelial neoplasia grade 3 or more severe (CIN3+) or CIN2+ diagnosed by the EP was compared between p16 IHC-negative and –positive CIN2 and between p16 IHC-positive CIN2 and CIN3 diagnosed by the CP using a Fisher's exact test. Finally, trends in HPV risk groups were compared for CIN2 diagnosed by the CP for all 4 combinations of cytology results (<HSIL vs. HSIL+) and p16 IHC results (negative and positive).

P values of <.05 were considered significant. STATA (Versions 13.1 and 15.1; StataCorp LLC, College Station, TX, USA) were used for analyses.

RESULTS

Correlations of HPV categories, p16 IHC results, and histologic diagnosis by the CP are shown in **Table 1**. The percent p16 IHC positive increased with an increasing severity of histologic diagnoses by the CP: 7.4% (95%CI = 4.7%-11.1%) for negatives, 26.6% (95%CI = 23.7%-29.5%) for CIN1, 71.9% (95%CI = 69.6%-74.2%) for CIN2, and 90.7% (95%CI = 88.9%-92.2%) for CIN3 (P_{trend}<.001); for reference, 94.5% (95%CI = 86.6%-98.5%) of the squamous cell carcinomas (SCC) tested p16-IHC positive. The percent p16-IHC positive increased with an increasing severity of biopsy diagnosis by the CP for each HPV risk group (i.e., HPV16 > HPV18/45 > other high-risk HPV > intermediate-risk HPV > low-risk HPV > HPV negative) (P_{trend}≤.001). The percent p16-IHC positive increased with an increasing severity of biopsy diagnosis by the CP even among HPV negatives (P_{trend}<.001), suggestive of some false-negative HPV genotyping. The percent p16-IHC positive increased with higher-risk HPV groups for each diagnosis (P_{trend}<.001). Testing p16 IHC positive was associated with higher HPV risk group than testing p16 IHC negative for each grade of CP-diagnosed biopsies (*P*<.001). That is, both HPV genotype and histologic diagnosis were independent determinants of testing p16 IHC positive.

We found no meaningful differences in the distribution of HPV genotypes detected or p16-IHC results based on the study sub-sample compared to those same results when extrapolated to the whole sample (**Supplemental Table 2**). HPV genotype-specific results stratified by the CP diagnosis and p16-IHC results are presented in **Supplemental Table 3**.

For all abnormal histology (CIN1 or more severe), p16 IHC-positive cases had higher-risk HPV than p16 IHC-negative diagnoses. Notably, p16 IHC-positive CIN2 had lower-risk

HPV than CIN3 (P_{trend}<.001); 38.2% (415 of 1,087) of p16 IHC-positive CIN2 tested positive for HPV16 whereas 54.5% (658 of 1,208) of all CIN3, regardless of the p16 IHC, tested positive for HPV16. p16 IHC-negative CIN2 had lower-risk HPV than p16 IHC-positive CIN2 (P_{trend}<.001) but higher-risk HPV than all CIN1 (P_{trend}<.001), p16 IHC positive CIN1 (P<.001), or negative histology (P_{trend}<.001) (i.e., regardless of the p16 IHC).

p16 IHC-positive CIN3 had higher-risk HPV than p16 IHC-negative CIN3 (P_{trend}<.001). p16 IHC-positive CIN3 was less likely to test positive for HPV18/45 (P<.001) than SCC, again demonstrating that HPV18/45 tends to under-represented in CIN3 compared to its attributable fraction in SCC.[3, 36, 37] p16 IHC-negative CIN3 had higher-risk HPV than p16 IHC-positive CIN2 (P_{trend}=.02). Similar patterns of p16 IHC and HPV risk groups were observed for the EP (**Supplemental Table 4**).

Table 2 shows the pair-wise diagnoses by the CP and EP and the correlation with testing p16 IHC positive or HPV16 positive, or having an antecedent HSIL+ cytology. Increasing severity of the EP diagnosis for a given CP diagnosis, and vice versa, was associated with increased likelihood of the biopsy testing p16-IHC positive or HPV16 positive (P_{trend}<.001 for both), with exception for when SCC was diagnosed by either the CP or the EP, or negative by the CP (P_{trend}<.001 for p16 IHC and P_{trend}=.40 for HPV16). Increasing severity of the CP diagnosis for a given EP diagnosis was also increasingly likely to have an antecedent HSIL+ cytology (P_{trend}<.001 for negative, CIN1, CIN2, and CIN3 and p=.05 for SCC). Increasing severity of the EP diagnosis for a CP diagnosis of CIN3 was more likely to have an antecedent HSIL+ cytology (P_{trend}<.001). Surprisingly, increasing severity of the EP diagnosis for a CP diagnosis of negative, CIN2, and cancer was not associated with having an antecedent HSIL+ cytology (P_{trend}>.05 for all).

Table 3 compares percent HPV16 positive, antecedent HSIL+ cytology, and CIN3+ and CIN2+ diagnosis by the EP between CP-diagnosed CIN2, stratified on p16 IHC status, and CIN3. p16 IHC-positive CIN2 was less likely than CIN3 to test HPV16 positive (38.18% vs. 54.47%, respectively, $P<.001$), have an antecedent HSIL+ cytology (21.02% vs. 42.53%, respectively, $P<.001$), or be diagnosed on review by the EP as CIN3+ (22.91% vs. 65.31%, respectively, $P<.001$) or CIN2+ (64.40% vs. 88.08%, respectively, $P<.001$). p16 IHC-positive CIN2 was less likely than CIN3 to be positive for at least one of these biomarkers (HPV16, antecedent HSIL+ cytology, and/or CIN3+ diagnosed by EP) (58.23% vs. 85.93%, respectively, $P<.001$). p16 IHC-positive CIN2 was less likely than CIN3 to be positive for all three biomarkers (2.48% vs. 15.89%, respectively, $P<.001$). p16 IHC-negative CIN2 was less likely than p16 IHC-positive CIN2 to be positive for any individual biomarker ($P<.001$), with exception of having antecedent HSIL+ cytology ($P>.99$). p16 IHC-negative CIN2 was less likely than p16 IHC-positive CIN2 to be positive for at least one biomarker ($P<.001$) or all three biomarkers ($P<.001$).

Table 4 compares the HPV risk group distribution for CP-diagnosed CIN2 that tested p16 IHC negative or positive and antecedent less than HSIL cytology (<HSIL) or HSIL+ cytology. Notably, p16 IHC-positive CIN2 with an antecedent HSIL+ cytology has lower risk HPV than all CP-diagnosed CIN3 ($P=.01$), p16 IHC-negative CIN2 with an antecedent <HSIL cytology has higher risk HPV than all CP-diagnosed CIN1 ($P<.001$).

DISCUSSION

In the largest case series to include HPV genotyping and p16 IHC immunostaining of biopsies to date, we were able to show the detailed relationship of these biomarkers with community diagnoses of precursors to cervical cancer, with a focus on CIN2. Key observations from our analyses were: 1) most CP-diagnosed CIN2 and CIN3, and a significant proportion of CIN1, tested p16-IHC positive; 2) p16 IHC-positive, CP-diagnosed CIN2 was less likely to test HPV16 positive, to have an antecedent HSIL+ cytology, and to be called CIN3+ or CIN2+ by the EP than CP-diagnosed CIN3; 3) p16 IHC-negative CIN2 had lower-risk HPV than p16 IHC-positive CIN2 but higher-risk HPV than CIN1; and 4) p16 IHC-negative CIN3 had higher-risk HPV than CIN2 or even p16 IHC-positive CIN2.

These data also confirm that p16 IHC corrects some of the errors in diagnosing high-grade cervical abnormalities but does so imperfectly. Approximately 28% of the CP-diagnosed CIN2 tested p16-IHC negative in this study; other studies have reported the percentage of p16 IHC-negative CIN2 ranging from approximately 20%[16] to less than 10%.[21, 38] Based on HPV risk group distribution, these cases were indeed lower risk and are less likely to progress to cancer. That is, p16 IHC-negative, CP-diagnosed CIN2 was more like CIN1 than CIN3. It is therefore justifiably to down-grade p16 IHC-negative, CP-diagnosed CIN2 to LSIL as recommended by LAST.[22] Conversely, p16 IHC-positive, CP-diagnosed CIN2 was more like CP-diagnosed CIN3 than CIN1.

However, these data also indicate that some fraction of the p16 IHC-positive, CP-diagnosed CIN2 is not a high-grade cervical abnormality. The implication of these data is that if LAST terminology[22] is to be used in routine practice (equivalent to CP), the use of HSIL

categorization for CIN3 or p16 IHC-positive CIN2 must include annotation of the H&E (morphologic) diagnosis, e.g., HSIL(CIN3) or HSIL(CIN2), respectively, which was suggested as optional by LAST. It is clear from these data that p16 IHC-positive CIN2 is NOT the clinical equivalent of CIN3. That is, a p16 IHC-positive CIN2 does not have the same clinical meaning (invasive potential) as CIN3, and therefore the two cannot be considered one clinical entity and should not be conflated with one another. When EP diagnosed the CP-diagnosed CIN2 biopsy as CIN3, the fraction that tested HPV16 and p16 IHC positive was close to that of the CP-diagnosed CIN3 biopsy, suggesting that these were high-grade cervical abnormalities. However, a consensus review by a panel of expert pathologists is not typically available in routine clinical practice.

Recent reports confirm the dramatic difference in risk of subsequent invasive cancer between CIN2[12,28] and CIN3[23] diagnoses. The disparity in HPV genotype distribution that we report provides a credible biologic explanation for this difference. Moreover, our findings emphasize that making this distinction between CIN2 and CIN3 for a HSIL diagnosis is necessary for clinical decision-making on whether to treat women with precursor lesions that might otherwise regress on their own. This is especially true in young women diagnosed with CIN2 for whom conservative management (wait and watch) is preferred[4] due possibly to the possibility of added risk of negative reproductive outcomes (e.g., preterm delivery) associated with excision treatments.[13, 14] Much of p16 IHC-positive CIN2 is likely to regress on its own, given that approximately 70-80% of CIN2 tests p16 IHC positive[16, 39], as observed here, but approximately 50% of all CIN2, and approximately 60% of CIN2 diagnosed in women under the age of 30 years, will regress.[12] Arithmetically, even if all p16 IHC-negative CIN2 was regressive, a significant proportion of p16 IHC-positive CIN2 must also be regressive. A prospective study

of women diagnosed with CIN2 reported that 57% of p16 IHC-positive CIN2 regressed in 12 months.[39] Retrospective study of women diagnosed with pathology review-confirmed CIN2 followed for two years found that 50% of regressive CIN2 were initially p16 IHC positive and 18% of p16 IHC-positive lesions regressed.[40]

In addition to causing the unnecessary treatment of some women with CIN2, losing the distinction between (p16-positive) CIN2 and CIN3 would also have long-term negative implications on the opportunity for future improvements to the diagnosis of high-grade cervical abnormalities. As new biomarkers are being developed that might better distinguish between HPV infection and high-grade cervical abnormalities and therefore might be applied to CIN2 or even p16 IHC-positive CIN2, it will be important to be able to easily identify such cases by qualifying whether they were diagnosed as CIN2 or CIN3.

These data also underscore the importance of the LAST recommendation *not* to perform p16 IHC testing systematically on all CIN3 or CIN1.[22] For CP-diagnosed CIN3, less than 10% of community diagnoses of CIN3 tested p16 IHC negative. CP-diagnosed CIN3 that tested p16 IHC negative had a less risky HPV group distribution ($P_{trend}<.001$) but were similarly likely to have antecedent high-grade cytology (25.5% vs. 24.8%, $P=.89$) as CP-diagnosed CIN2 that was diagnosed as CIN3 by the EP. Thus, p16 IHC-negative CIN3 is unlikely to be at sufficiently low risk to change its management i.e., there is no clinical utility, only added cost, to systematically using p16-IHC on CIN3 diagnoses.

Nor is there evidence that p16 IHC testing of CIN1 provides clinically meaningful risk stratification or predicts progression to CIN2+ as previously shown.[41-43] Here, based on HPV risk group distribution, p16 IHC-positive CIN1 was higher risk than p16 IHC-negative CIN1 but not

even as high risk as p16 IHC-negative CIN2. Moreover, p16 IHC-positive CIN1 was similarly unlikely to have an antecedent high-grade cytology as p16 IHC-negative CIN1 (3.2% vs. 5.1%, respectively, P=.27) (data not shown).

Aside from the added cost for limited or no benefit to women diagnosed with CIN1, p16-IHC testing of biopsies diagnosed as CIN1 might result in incorrect, over-interpretation of a positive p16-IHC result as CIN2, which could then lead to unnecessary treatment and a concomitant increased risk of preterm delivery for those still considering childbearing.[13, 14, 42]

Limited p16 IHC testing, and/or possibly Ki-67 IHC testing, of some CIN1 might have some value for internal use as a laboratory quality control standard[7, 44], similar to the use of HPV:SIL rates for cytology[45], to set the threshold of normal vs. non-normal histology. We observed that approximately one-quarter of the CP-diagnosed CIN1 tested p16-IHC positive in this study. Other studies have reported a percent p16-IHC positive for CIN1 ranging from approximately 10% to almost 60%[16, 21, 42, 46-50], suggesting significant variability/unreliability in the morphologic interpretation of diagnosis criteria for CIN1 (vs. negative or CIN2) compared to a more objective standard i.e., p16 IHC.

Likewise, some pathologists equivocate between CIN1 and CIN2 or diagnosis “CIN1/2”. We did not separately analyze cases of CIN1/2 as there were small numbers in our dataset. However, the LAST recommendation for using p16 IHC was to clarify the clinical meaning of CIN2 by distinguish those CIN2 that were higher risk (p16 IHC-positive CIN2) from those that were lower risk (p16 IHC-negative CIN2), not to clarify meaning of CIN1. Here again, there would be potential for over-utilization of p16 IHC. Pathologists uncertain whether a biopsy diagnosis is CIN1 or CIN2 might be tempted to call it CIN1/2 or even CIN2 for perceived

greater safety, believing that using p16 IHC as an adjunctive test would correct any overcalls. However, because such a high percentage of CIN1 (and therefore also "CIN1/2") will also test p16 IHC positive, most of which is nothing more than low-grade, benign, regressive CIN1, these biopsies might then get misclassified as HSIL and women would receive unnecessary treatment.

A few other scenarios might exacerbate the inappropriate use of p16 IHC on CIN1 biopsies. First, pathologists, worried that future review of a CIN1 biopsy by second pathologist might result in a CIN2 diagnosis, might be motivated to do p16 IHC on a CIN1 or a CIN1/2 that previously they would report as CIN1. In addition, p16 IHC on CIN1 or CIN1/2 may be used by pathologists as feedback to lower their criteria for a CIN2 diagnosis resulting in more CIN1 being called CIN2.

We acknowledge an important limitation: this analysis was cross-sectional and therefore we could only make inferences related to true cervical cancer risk based on biomarker distributions. Nevertheless, these biomarkers included in this analysis are strong predictors of cervical cancer risk and if practical, could be incorporated into improved diagnostic classification of cervical abnormalities. Indeed, women whose Pap specimen was called high-grade cytology and tested HPV16 positive (the read out for which is provided by some HPV tests vs. setting up laboratory testing of biopsies) are at very high risk of CIN3, up to ~80%.[51, 52]

Another potential limitation is that the EP diagnoses were based on H&E alone and were not informed by p16 IHC results. However, it is unknown whether p16-IHC informed interpretations of H&E diagnoses would have resulted in improved classification of CIN2 vs. H&E diagnosis rendered independently of p16 IHC results.

Future studies in cohorts with long-term follow-up will be needed to determine the risk stratification provided by p16-IHC testing of CIN2. These cohorts will need to be rather large because of the rarity of CIN2 (<1%) in the general population and losses to follow-up. Alternatively, retrospective analyses of conservatively managed CIN2 in younger women could be done in which the index CIN2 biopsies are tested by p16 IHC. Such studies would provide important information on its risk stratification, how safe (vs. invasive cancer) women with p16 IHC-negative CIN2 are, and how much over-treatment is likely to occur if p16 IHC-positive CIN2 were to be treated immediately.

It is clear from these and other data that p16 IHC is a sensitive but non-specific biomarker of CIN3, which is a more rigorous definition of "cervical precancer" than CIN2 and even p16 IHC-positive CIN2. Even so, many but not all CIN3 will develop into invasive cervical cancer if left untreated.[23] However, because of the lack of specificity of p16 IHC, presumably due to its increased expression in response to productive HPV infections that may or may not progress[53, 54], many low-grade cervical abnormalities will still test p16-IHC positive even if they are not destined to progress to high-grade cervical abnormalities. LAST classification of "HSIL" diagnosis, which includes p16 IHC-positive CIN2, should annotate the morphologic diagnosis of CIN2 or CIN3 in routine clinical practice to inform all clinical management decisions. This is especially important for (but not limited to) those women under the age of 30 years and/or considering childbearing and diagnosed with CIN2 for whom surveillance rather than treatment is recommended[4] and/or desirable, respectively. While neither "biomarker", CIN diagnosis or p16 IHC, is perfect, together they further stratify the cervical cancer risk and if used correctly better inform clinical decision making than either can accomplish alone.

In summary, there is currently no reliable way to distinguish those CIN2 and even p16 IHC-positive CIN2 diagnoses that will progress or regress. The use of other biomarkers with or without p16 IHC may improve the diagnostic classification of cervical abnormalities in relation to their invasive cancer potential.

Acknowledgements: The views expressed in this publication are those of the authors alone and do reflect official policy of the Department of the Army, the Department of Defense, the Centers for Disease Control and Prevention, the National Institutes of Health, or the United States Government.

Funding: Information reported in this publication was supported by the National Institute of Allergy and Infectious Diseases (NIAID) of the US National Institutes of Health under award number U19AI113187 and through its Developmental Research Pilot (DRP) program. Roche Molecular Systems and Ventana, a Roche Company, provided reagents and equipment at no cost.

New Mexico HPV Pap Registry (NMHPVPR) Steering Committee Members: Members of the New Mexico HPV Pap Registry (NMHPVPR) Steering Committee reviewed and gave input to the manuscript and supported the concept and directions of the NMHPVPR including the evaluations presented in this manuscript. The NMHPVPR Steering members participating in this effort are as follows: Nancy E. Joste, MD, University of New Mexico Health Sciences Center and Tricore Reference Laboratories, Albuquerque, New Mexico; Walter Kinney, MD, Kaiser Permanente Northern California; Cosette M. Wheeler, PhD, University of New Mexico Health Sciences Center; William C. Hunt, MS, University of New Mexico Health Sciences Center; Alan Waxman, MD MPH, University of New Mexico Health Sciences Center; Steven Jenison, MD, Community Member; Julia C. Gage, PhD, MPH, US National Cancer Institute; Mark Schiffman, MD, MPH, US National Cancer Institute; Philip E. Castle, PhD MPH, Albert Einstein School of Medicine; Vicki Benard, PhD, US Centers for Disease Control and Prevention; Debbie Saslow, PhD,

American Cancer Society; Jane J. Kim PhD, Harvard TH Chan School of Public Health; Mark H. Stoler MD, University of Virginia; Jack Cuzick, PhD, Wolfson Institute of Preventive Medicine, London; Patti Gravitt, PhD, University of New Mexico; Giovanna Rossi Pressley, MSc, Collective Action Strategies; and Kevin English, DrPh MPH, Albuquerque Area Southwest Tribal Epidemiology Center (AASTEC). No compensation was received for contributions to this manuscript by any named authors or by the NMHPVPR Steering Committee members.

The p16 IHC Study Panel (SP) Group Members: James H. Bainton, MD, Kaiser South San Francisco Medical Center, San Francisco, CA; Sarah M. Bean, MD, Duke University, Durham, NC; Michael J. Beckmann, DO, Southern Illinois University School of Medicine, Springfield, IL; Cynthia Benedict, MD, DCL Pathology, Indianapolis, IN; Sarah Cherny, MD, Kaiser South San Francisco Medical Center, San Francisco, CA; Vincenzo Ciocca, MD, Lankenau Medical Center, Wynnewood, PA; Barbara A. Crothers COL MC USA, Walter Reed National Military Medical Center, Bethesda, MD; Caishu Deng, MD, TPMG Modesto Medical Center, Modesto, CA; Marie Dvorakova, MD, University of Calgary, Calgary, Alberta; Fang Fan, MD, University of Kansas Medical Center, Kansas City, KS; Mohiedean Ghofrani, MD, MBA, Peace Health Laboratories, Vancouver, WA; Krisztina Hanley, MD, Emory, University Hospital, Atlanta, GA; Debra S. Heller, MD, Rutgers New Jersey Medical School, Newark, NJ; Mujtaba Husain, MD, University of Central Florida, Orlando, FL; Kristin Jensen, MD, Veterans Affairs Palo Alto Health Care System and Stanford University Hospital, Stanford, CA; Spencer Kerley, MD, Physicians Reference Laboratory, Overland Park, KS; Linda Korfhage, MD, Norton Audubon Hospital, Louisville, KY; Kruti P. Maniar, MD, Northwestern University, Chicago, IL; Carrie Marshall, MD, University of Colorado Hospital, Aurora, CO; Karen Matsukuma, MD, University of California Davis, Sacramento, CA; Zulfia McCroskey, MD, Loyola University Medical Center, Maywood, IL; Grace Moore, MD, CPA Lab, Louisville, KY; Jeannie Muir, LTC, MC, Walter Reed National Medical Center, Bethesda, MD; Michael Nowacki, MD, Norton

Hospital Laboratory, Louisville, KY; Idris Tolgay Ocal, MD, Mayo Clinic Arizona, Scottsdale, AZ; Dennis O'Connor, MD, CPA Labs, Louisville, KY; Lori Scanga, MD, Women and Children's Hospitals, Chapel Hill, NC; Malcom Schinstine, PhD, MD, Diagnostic and Research Centre, Calgary, AB; Peter Schlosshauer, MD, Sunrise Medical Laboratories, Hicksville, NY; Charles Sturgis, MD, Cleveland Clinic, Cleveland, OH; Jean Thomsen, MD, Jennie Edmundson Hospital, Omaha, NE; Diane Tokugawa, MD, TPMG Regional Laboratory, Berkeley, CA; Elizabeth Unger, MD, Centers for Disease Control and Prevention, Atlanta, GA; Keith Volmar, MD, Rex Pathology Associates, Raleigh, NC; Yinong Wang, MD, Clinical Laboratory Services, Calgary, AB; Dava West, MD, University of Kentucky, Lexington, KY; Sandra White, MD, Puget Sound Institute of Pathology, Seattle, WA; Barbara Winkler, MD, Mount Kisco Medical Group, Mount Kisco, NY; Anna Yemelyanova, MD, UT MD Anderson Cancer Center, Houston, TX; Chengquan Zhao, MD, Magee Women's Hospital, Pittsburgh, PA.

Table 1. Human papillomavirus (HPV) genotyping results, categorized according to cervical-cancer risk, and p16 immunohistochemistry (IHC) results (p16 IHC positive [p16+] or negative [p16-]) for biopsy diagnosis by the community pathology (CP) biopsy diagnoses of negative, CIN1, CIN2, and CIN3. Results for all squamous cell carcinoma (SCC) are shown for reference. "%Col" is the column percentage i.e., the number in cell divide by the total column number.

Diagnosis:	Negative				CIN1				CIN2				CIN3					SCC	
	p16-IHC Negative		p16-IHC Positive		p16-IHC Negative		p16-IHC Positive		p16-IHC Negative		p16-IHC Positive		p16-IHC Negative		p16-IHC Positive			All	
HPV Categories‡	**N**	**%Col**	**N**	**%Col**	**N**	**%Col**	**N**	**%Col**	**N**	**%Col**	**N**	**%Col**	**N**	**%Col**	**N**	**%Col**	p_{trend}†	**N**	**%Col**
HPV16	8	3.1	1	4.8	32	4.7	35	14.1	91	21.4	415	38.2	41	36.3	617	56.3	<.001	43	58.9
HPV18/45	5	1.9	2	9.5	24	3.5	25	10.1	29	6.8	82	7.5	8	7.1	57	5.2	<.001	10	13.7
Other High-Risk*	21	8.0	7	33.3	142	20.7	128	51.6	163	38.4	496	45.6	36	31.9	354	32.3	<.001	10	13.7
Intermediate Risk**	9	3.4	2	9.5	54	7.9	28	11.3	29	6.8	56	5.2	3	2.7	31	2.8	<.001	2	2.7
Low-Risk¥	4	1.5	0	0.0	35	5.1	4	1.6	18	4.2	7	0.6	4	3.5	3	0.3	.0097	1	1.4
HPV Negative	215	82.1	9	42.9	399	58.2	28	11.3	95	22.4	31	2.9	21	18.6	33	3.0	<.001	7	9.6
Total	262	100	21	100	686	100	248	100	425	100	1,087	100	113	100	1,095	100	<.001	73	100
p_{trend}§ =	<.001				<.001				<.001				<.001						

‡Defined hierarchically according to cancer risk

†test of trend for testing p16 IHC positive across diagnoses (excluding cancer) for each HPV risk group

*HPV31, 33, 35, 39, 51, 52, 56, 58, 59, and 68

**HPV26, 53, 66, 67, 70, 73, and 82

¥HPV6, 11, 40, 42, 54, 55, 61, 62, 64, 69, 71, 72, 81, 82v, 83, 84, and 89

§test of trend for HPV risk group by p16 IHC result for each diagnosis

Table 2. Pairwise diagnoses by the community pathology (CP) and expert panel (EP) and the percent p16 immunohistochemistry (IHC) positive (%p16 IHC+), human papillomavirus type 16 (HPV16) positive (HPV16+), and with an antecedent high-grade squamous intraepithelial lesion (HSIL) or more severe (HSIL+) cytologic interpretation. "%cell" is the cell percentage i.e., N_{cell}/N_{total}.

			Expert Panel (EP) Diagnosis								
			Negative	CIN1	CIN2	CIN3	SCC	Total	p_{trend} (p16)	p_{trend} (HPV16)	p_{trend} (HSIL+ Cytology)
Community Pathology (CP) Diagnosis	Negative	**N**	**261**	**14**	**7**	**1**	**0**	283	.009	.40	.15
		%cell	*6.53*	*0.35*	*0.18*	*0.03*	*0.00*	*7.08*			
		%p16 IHC+	6.51	7.14	42.86	0.00	0.00	7.42			
		%HPV16+	3.07	0.00	14.29	0.00	0.00	3.18			
		%HSIL+ Cytology*	7.58	33.33	14.29	0.00	0.00	8.84			
	CIN1	**N**	**535**	**318**	**67**	**11**	**0**	931	<.001	<.001	.029
		%cell	*13.39*	*7.96*	*1.68*	*0.28*	*0.00*	*23.29*			
		%p16 IHC+	10.47	42.77	68.66	81.82	0.00	26.53			
		%HPV16+	3.74	11.32	16.42	0.00	0.00	7.20			
		%HSIL+ Cytology**	5.93	3.25	1.56	0.00	0.00	4.62			
	CIN2	**N**	**260**	**431**	**547**	**271**	**0**	1509	<.001	<.001	.43
		%cell	*6.50*	*10.78*	*13.69*	*6.78*	*0.00*	*37.75*			
		%p16 IHC+	34.23	68.68	82.45	91.88	0.00	71.90			
		%HPV16+	18.46	24.59	40.04	47.97	0.00	33.33			
		%HSIL+ Cytology§	23.45	17.66	20.80	24.80	0.00	21.08			
	CIN3	**N**	**90**	**48**	**275**	**787**	**2**	1202	<.001	<.001	<.001
		%cell	*2.25*	*1.20*	*6.88*	*19.69*	*0.05*	*33.07*			
		%p16 IHC+	42.22	70.83	89.45	97.71	100.00	90.60			
		%HPV16+	33.33	25.00	49.45	60.36	50.00	54.41			
		%HSIL+ Cytology†	27.50	31.82	34.60	47.74	0.00	42.50			
	SCC	**N**	**1**	**0**	**2**	**17**	**52**	72	.010	.36	.17
		%cell	*0.03*	*0.00*	*0.05*	*0.43*	*1.30*	*1.80*			
		%p16 IHC+	0.00	0.00	100.00	100.00	96.15	95.83			
		%HPV16+	0.00	0.00	50.00	58.82	59.62	58.33			
		%HSIL+ Cytology‡	0.00	0.00	100.00	75.0	82.76	79.55			
	Total	**N**	1147	811	898	1087	54	3997			
		%cell	*28.70*	*20.29*	*22.47*	*27.20*	*1.35*	10.00			
		%p16 IHC+	17.44	57.58	83.30	96.04	96.30	62.82			
		%HPV16+	9.24	18.99	40.98	56.58	59.26	31.90			

		%HSIL+ Cytology‡	12.08	13.15	23.46	41.61	80.00	23.61
	p_{trend} (p16) =		<.001	<.001	<.001	<.001	.78	<.001
	p_{trend} (HPV16) =		<.001	<.001	<.001	<.001	.79	<.001
	p_{trend} (HSIL+ Cytology) =		<.001	<.001	<.001	<.001	.046	<.001

13 cases excluded as diagnosed as adenocarcinoma *in situ* or adenocarcinoma by EP (3 CP [redacted] CP SCC)

*68 Missing Cytology (63 EP negative, 5 EP CIN1)

**108 Missing Cytology (63 EP negative, 41 EP CIN1, 3 EP CIN2, 1 EP CIN3)

§152 Missing Cytology (34 EP negative, 46 EP CIN1, 47 EP CIN2, 25 EP CIN3)

†155 Missing Cytology (10 EP negative, 4 EP CIN1, 38 EP CIN2, 102 EP CIN3, 1 EP SCC)

‡28 Missing Cytology (5 EP CIN3, 23 EP SCC [squamous cell carcinoma])

Table 3. The relationships of community pathology-diagnosed CIN3 and CIN2, stratified on p16 IHC result, with biomarkers of cervical cancer risk: the biopsy testing positive for human papillomavirus type 16 (HPV16), an antecedent high-grade intraepithelial lesion (HSIL) or more severe (HSIL+) cytologic interpretation, and an expert panel review histopathological diagnosis of cervical intraepithelial neoplasia (CIN) grade 3 (CIN3) or more severe (CIN3+) or CIN grade 2 (CIN2) or more severe (CIN2+). Below the individual biomarker results, the relationships of the diagnoses with combinations of any (or) or all (and) biomarkers are shown.

	p16 IHC-Negative CIN2		**p16 IHC-Positive CIN2**		**CIN3**		**p**[†]	**p$_{trend}$**[‡]
Biomarker Result	**N**	**%**	**N**	**%**	**N**	**%**		
HPV16 Positive	91	21.41	415	38.18	658	54.47	<.001	<.001
HSIL+ Cytology*	80	21.11	206	21.02	447	42.53	<.001	<.001
EP Diagnosis of CIN3+	22	5.18	249	22.91	789	65.31	<.001	<.001
EP Diagnosis of CIN2+	118	27.76	700	64.40	1,064	88.08	<.001	<.001
HPV16 Positive, HSIL+, and/or EP Diagnosis of CIN3+	163	38.35	633	58.23	1,038	85.93	<.001	<.001
HPV16 Positive, HSIL+, and EP Diagnosis of CIN3+	3	0.71	27	2.48	192	15.89	<.001	<.001

*****46 p16 IHC-Negative CIN2, 107 p16 IHC-Positive CIN2 and 157 CIN3 missing antecedent cytology

[†]p16 IHC-positive CIN2 vs. CIN3

[‡]trend for p16 IHC-negative CIN2 vs. p16 IHC-positive CIN2 vs. CIN3

Table 4. The relationships of community pathology (CP)-diagnosed cervical intraepithelial neoplasia grade 2 (CIN2), stratified by p16 immunohistochemistry (IHC) results and antecedent cytologic interpretation categorized as high-grade squamous intraepithelial lesion (HSIL) or more severe (HSIL+) vs. not (<HSIL), with human papillomavirus (HPV) categories and compared to CP-diagnosed CIN3. One hundred fifty-three cases were missing antecedent cytology results. "%Col" is the column percentage i.e., the number in cell divide by the total column number.

	Community Pathology-Diagnosed CIN2							
	p16 IHC Negative		**p16 IHC Negative**		**p16 IHC Positive**		**p16 IHC Positive**	
	<HSIL Cytology		**HSIL+ Cytology**		**<HSIL Cytology**		**HSIL+ Cytology**	
HPV Risk Group‡	**N**	**%col**	**N**	**%col**	**N**	**%col**	**N**	**%col**
HPV16	61	20.4	18	22.5	284	36.7	88	42.7
HPV18/45	16	5.4	9	11.3	54	7.0	17	8.3
Other High Risk*	127	42.5	28	35.0	360	46.5	87	42.2
Intermediate Risk**	21	7.0	5	6.3	46	5.9	8	3.9
Low Risk¥	15	5.0	1	1.3	6	0.8	0	0.0
HPV Negative	59	19.7	19	23.8	24	3.1	6	2.9
p_{trend} (vs. CIN3†**)**	<.001		<.001		<.001		.012	

‡Defined hierarchically according to cancer risk

*HPV31, 33, 35, 39, 51, 52, 56, 58, 59, and 68

**HPV26, 53, 66, 67, 70, 73, and 82

¥HPV6, 11, 40, 42, 54, 55, 61, 62, 64, 69, 71, 72, 81, 82v, 83, 84, and 89

†Compared to data combining p16 IHC-negative and p16 IHC-positive CIN3 from **Table 1**

Figure 1. Consort diagram of specimen inclusions and exclusions. Abbreviations: CIN, cervical intraepithelial neoplasia; CIN2, CIN grade 2; CIN3, CIN grade 3; AIS, adenocarcinoma *in situ*; ADCA, adenocarcinoma; SCC, squamous cell carcinoma; CP, community pathologists; EP, expert pathologists

Reference List

(1) Walboomers JM, Jacobs MV, Manos MM et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol* 1999;189(1):12-19.

(2) de Sanjose S, Quint WG, Alemany L et al. Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study. *Lancet Oncol* 2010;11(11):1048-1056.

(3) Guan P, Howell-Jones R, Li N et al. Human papillomavirus types in 115,789 HPV-positive women: a meta-analysis from cervical infection to cancer. *Int J Cancer* 2012;131(10):2349-2359.

(4) Massad LS, Einstein MH, Huh WK et al. 2012 updated consensus guidelines for the management of abnormal cervical cancer screening tests and cancer precursors. *J Low Genit Tract Dis* 2013;17(5 Suppl 1):S1-S27.

(5) Schiffman M, Castle PE, Jeronimo J, Rodriguez AC, Wacholder S. Human papillomavirus and cervical cancer. *Lancet* 2007 ;370(9590):890-907.

(6) Richart RM. A theory of cervical carcinogenesis. *Obstet Gynecol Surv* 1969;24(7 Pt 2):874-879.

(7) van Zummeren M, Leeman A, Kremer WW et al. Three-tiered score for Ki-67 and p16(ink4a) improves accuracy and reproducibility of grading CIN lesions. *J Clin Pathol* 2018;71(11):981-988.

(8) Stoler MH, Schiffman M. Interobserver reproducibility of cervical cytologic and histologic interpretations: realistic estimates from the ASCUS-LSIL Triage Study. *JAMA* 2001;285(11):1500-1505.

(9) Carreon JD, Sherman ME, Guillen D et al. CIN2 is a much less reproducible and less valid diagnosis than CIN3: results from a histological review of population-based cervical samples. *Int J Gynecol Pathol* 2007;26(4):441-446.

(10) Stoler MH, Ronnett BM, Joste NE, Hunt WC, Cuzick J, Wheeler CM. The interpretive variability of cervical biopsies and its relationship to HPV status. *Am J Surg Pathol* 2015;39(6):729-736.

(11) Reuschenbach M, Wentzensen N, Dijkstra MG, von Knebel DM, Arbyn M. p16INK4a Immunohistochemistry in Cervical Biopsy Specimens: A Systematic Review and Meta-Analysis of the Interobserver Agreement. *Am J Clin Pathol* 2014;142(6):767-772.

(12) Tainio K, Athanasiou A, Tikkinen KAO et al. Clinical course of untreated cervical intraepithelial neoplasia grade 2 under active surveillance: systematic review and meta-analysis. *BMJ* 2018;360:k499.:k499.

(13) Sasieni P, Castanon A, Landy R et al. Risk of preterm birth following surgical treatment for cervical disease: executive summary of a recent symposium. *BJOG* 2016;123(9):1426-1429.

(14) Kyrgiou M, Athanasiou A, Paraskevaidi M et al. Adverse obstetric outcomes after local treatment for cervical preinvasive and early invasive disease according to cone depth: systematic review and meta-analysis. *BMJ* 2016;354:i3633. doi: 10.1136/bmj.i3633.:i3633.

(15) Galgano MT, Castle PE, Stoler MH, Solomon D, Schiffman M. Can HPV-16 genotyping provide a benchmark for cervical biopsy specimen interpretation? *Am J Clin Pathol* 2008;130(1):65-70.

(16) Galgano MT, Castle PE, Atkins KA, Brix WK, Nassau SR, Stoler MH. Using biomarkers as objective standards in the diagnosis of cervical biopsies. *Am J Surg Pathol* 2010;34(8):1077-1087.

(17) Yemelyanova A, Gravitt PE, Ronnett BM et al. Immunohistochemical detection of human papillomavirus capsid proteins L1 and L2 in squamous intraepithelial lesions: potential utility in diagnosis and management. *Mod Pathol* 2013;26(2):268-274.

(18) Griffin H, Soneji Y, van Baars R et al. Stratification of HPV-induced cervical pathology using the virally encoded molecular marker E4 in combination with p16 or MCM. *Mod Pathol* 2015;28(7):977-993.

(19) van Baars R, Griffin H, Wu Z et al. Investigating diagnostic problems of CIN1 and CIN2 associated with high-risk HPV by combining the novel molecular biomarker PanHPVE4 With p16INK4a. *Am J Surg Pathol* 2015;39(11):1518-1528.

(20) Klaes R, Benner A, Friedrich T et al. p16INK4a immunohistochemistry improves interobserver agreement in the diagnosis of cervical intraepithelial neoplasia. *Am J Surg Pathol* 2002;26(11):1389-1399.

(21) Stoler MH, Wright TC, Jr., Ferenczy A et al. Routine use of adjunctive p16 immunohistochemistry improves diagnostic agreement of cervical biopsy interpretation: Results from the CERTAIN study. *Am J Surg Pathol* 2018;42(8):1001-1009.

(22) Darragh TM, Colgan TJ, Cox JT et al. The Lower Anogenital Squamous Terminology Standardization Project for HPV-Associated Lesions: background and consensus recommendations from the College of American Pathologists and the American Society for Colposcopy and Cervical Pathology. *Arch Pathol Lab Med* 2012;136(10):1266-1297.

(23) McCredie MR, Sharples KJ, Paul C et al. Natural history of cervical neoplasia and risk of invasive cancer in women with cervical intraepithelial neoplasia 3: a retrospective cohort study. *Lancet Oncol* 2008;9(5):425-434.

(24) Castle PE, Schiffman M, Wheeler CM, Wentzensen N, Gravitt PE. Impact of improved classification on the association of human papillomavirus with cervical precancer. *Am J Epidemiol* 2010;171(2):155-163.

(25) Castle PE, Stoler MH, Solomon D, Schiffman M. The relationship of community biopsy-diagnosed cervical intraepithelial neoplasia grade 2 to the quality control pathology-reviewed diagnoses: An ALTS report. *Am J Clin Pathol* 2007;127(5):805-815.

(26) Hosaka M, Fujita H, Hanley SJ et al. Incidence risk of cervical intraepithelial neoplasia 3 or more severe lesions is a function of human papillomavirus genotypes and severity of cytological and histological abnormalities in adult Japanese women. *Int J Cancer* 2013;132(2):327-334.

(27) Silver MI, Gage JC, Schiffman M et al. Clinical outcomes after conservative management of cervical intraepithelial neoplasia grade 2 (CIN2) in women ages 21-39 years. *Cancer Prev Res (Phila)* 2018;11(3):165-170.

(28) Joste NE, Ronnett BM, Hunt WC et al. Human papillomavirus genotype-specific prevalence across the continuum of cervical neoplasia and cancer. *Cancer Epidemiol Biomarkers Prev* 2015;24(1):230-240.

(29) Gravitt PE, Peyton CL, Alessi TQ et al. Improved amplification of genital human papillomaviruses. *J Clin Microbiol* 2000;38(1):357-361.

(30) Peyton CL, Gravitt PE, Hunt WC et al. Determinants of genital human papillomavirus detection in a US population. *J Infect Dis* 2001;183(11):1554-1564.

(31) Gravitt PE, Schiffman M, Solomon D, Wheeler CM, Castle PE. A comparison of linear array and hybrid capture 2 for detection of carcinogenic human papillomavirus and cervical precancer in ASCUS-LSIL triage study. *Cancer Epidemiol Biomarkers Prev* 2008;17(5):1248-1254.

(32) Wheeler CM, Hunt WC, Cuzick J et al. A population-based study of human papillomavirus genotype prevalence in the United States: baseline measures prior to mass human papillomavirus vaccination. *Int J Cancer* 2013;132(1):198-207.

(33) Munoz N, Bosch FX, de Sanjose S et al. Epidemiologic classification of human papillomavirus types associated with cervical cancer. *N Engl J Med* 2003;348(6):518-527.

(34) Cuzick J. A Wilcoxon-type test for trend. *Stat Med* 1985;4(1):87-90.

(35) Solomon D, Davey D, Kurman R et al. The 2001 Bethesda System: terminology for reporting results of cervical cytology. *JAMA* 2002;287(16):2114-2119.

(36) Wheeler CM, Hunt WC, Joste NE, Key CR, Quint WG, Castle PE. Human papillomavirus genotype distributions: implications for vaccination and cancer screening in the United States. *J Natl Cancer Inst* 2009;101(7):475-487.

(37) Safaeian M, Schiffman M, Gage J, Solomon D, Wheeler CM, Castle PE. Detection of precancerous cervical lesions is differential by human papillomavirus type. *Cancer Res* 2009;69(8):3262-3266.

(38) Arvizo C, Chen Q, Du H et al. p16 immunohistochemistry in colposcope-directed and random cervical biopsies of CIN2 and CIN3. *J Low Genit Tract Dis* 2016;20(3):197-200.

(39) Miralpeix E, Genoves J, Maria Sole-Sedeno J et al. Usefulness of p16(INK4a) staining for managing histological high-grade squamous intraepithelial cervical lesions. *Mod Pathol* 2017;30(2):304-310.

(40) Miyamoto S, Hasegawa J, Morioka M, Hirota Y, Kushima M, Sekizawa A. The association between p16 and Ki-67 immunohistostaining and the progression of cervical intraepithelial neoplasia grade 2. *Int J Gynaecol Obstet* 2016;134(1):45-48.

(41) Mills AM, Paquette C, Terzic T, Castle PE, Stoler MH. CK7 immunohistochemistry as a predictor of CIN1 progression: a retrospective study of patients from the quadrivalent HPV vaccine trials. *Am J Surg Pathol* 2017;41(2):143-152.

(42) Pacchiarotti A, Ferrari F, Bellardini P et al. Prognostic value of p16-INK4A protein in women with negative or CIN1 histology result: a follow-up study. *Int J Cancer* 2014;134(4):897-904.

(43) Sagasta A, Castillo P, Saco A et al. p16 staining has limited value in predicting the outcome of histological low-grade squamous intraepithelial lesions of the cervix. *Mod Pathol* 2016;29(1):51-59.

(44) Singh C, Manivel JC, Truskinovsky AM et al. Variability of pathologists' utilization of p16 and ki-67 immunostaining in the diagnosis of cervical biopsies in routine pathology practice and its impact on the frequencies of cervical intraepithelial neoplasia diagnoses and cytohistologic correlations. *Arch Pathol Lab Med* 2014;138(1):76-87.

(45) Sherman ME, Schiffman MH, Lorincz AT et al. Toward objective quality assurance in cervical cytopathology. Correlation of cytopathologic diagnoses with detection of high-risk human papillomavirus types. *Am J Clin Pathol* 1994;102(2):182-187.

(46) Zhong P, Li J, Gu Y et al. P16 and Ki-67 expression improves the diagnostic accuracy of cervical lesions but not predict persistent high risk human papillomavirus infection with CIN1. *Int J Clin Exp Pathol* 2015 ;8(3):2979-2986.

(47) Selvi K, Badhe BA, Papa D, Ganesh RN. Role of p16, CK17, p63, and human papillomavirus in diagnosis of cervical intraepithelial neoplasia and distinction from its mimics. *Int J Surg Pathol* 2014;22(3):221-230.

(48) Jackson JA, Kapur U, Ersahin C. Utility of p16, Ki-67, and HPV test in diagnosis of cervical intraepithelial neoplasia and atrophy in women older than 50 years with 3- to 7-year follow-up. *Int J Surg Pathol* 2012;20(2):146-153.

(49) Guo M, Baruch AC, Silva EG et al. Efficacy of p16 and ProExC immunostaining in the detection of high-grade cervical intraepithelial neoplasia and cervical carcinoma. *Am J Clin Pathol* 2011;135(2):212-220.

(50) Razmpoosh M, Sansregret A, Oligny LL et al. Assessment of correlation between p16INK4a staining, specific subtype of human papillomavirus, and progression of LSIL/CIN1 lesions: first comparative study. *Am J Clin Pathol* 2014;142(1):104-110.

(51) Silver MI, Andrews J, Cooper CK et al. Risk of cervical intraepithelial neoplasia 2 or worse by cytology, human papillomavirus 16/18, and colposcopy impression: a systematic review and meta-analysis. *Obstet Gynecol* 2018;132(3):725-735.

(52) Castle PE, Aslam S, Behrens C. Cervical precancer and cancer risk by human papillomavirus status and cytologic interpretation: implications for risk-based management. *Cancer Epidemiol Biomarkers Prev* 2016;25(12):1595-1599.

(53) von Knebel DM, Reuschenbach M, Schmidt D, Bergeron C. Biomarkers for cervical cancer screening: the role of p16(INK4a) to highlight transforming HPV infections. *Expert Rev Proteomics* 2012;9(2):149-163.

(54) Sahasrabuddhe VV, Luhn P, Wentzensen N. Human papillomavirus and cervical cancer: biomarkers for improved prevention efforts. *Future Microbiol* 2011;6(9):1083-1098.